BMC Biotechnology

This Provisional PDF corresponds to the article as it appeared upon acceptance. The fully-formatted PDF version will become available in approximately two weeks after the date of publication, from the URL listed below.

Performance of high-throughput DNA quantification methods

Kashif A Haque (haquek@mail.nih.gov)
Ruth M Pfeiffer (pfeiffer@mail.nih.gov)
Michael B Beerman (beerman@mail.nih.gov)
Jeff P Struewing (struewij@mail.nih.gov)
Stephen J Chanock (chanocks@mail.nih.gov)
Andrew W Bergen (bergena@mail.nih.gov)

Performance of high-throughput DNA quantification methods

Kashif A Haque[1,*], Ruth M Pfeiffer[2], Michael B Beerman[1], Jeff P Struewing[3], Stephen J Chanock[4,] Andrew W Bergen[5]

1 Core Genotyping Facility, SAIC-Frederick, Inc., National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA.
2 Biostatistics Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA.
3 Laboratory of Population Genetics, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA.
4 Section on Genomic Variation, Pediatric Oncology Branch, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD 20879, USA.
5 Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA.

* To whom correspondence should be addressed.

Abstract

Background

The accuracy and precision of estimates of DNA concentration are critical factors for efficient use of DNA samples in high-throughput genotype and sequence analyses. We evaluated the performance of spectrophotometric (OD) DNA quantification, and compared it to two fluorometric quantification methods, the PicoGreen® assay (PG), and a novel real-time quantitative genomic PCR assay (QG) specific to a region at the human *BRCA1* locus. Twenty -Two lymphoblastoid cell line DNA samples with an initial concentration of ~350 ng/uL were diluted to 20 ng/uL. DNA concentration was estimated by OD and further diluted to 5 ng/uL. The concentrations of multiple aliquots of the final dilution were measured by the OD, QG and PG methods. The effects of manual and robotic laboratory sample handling procedures on the estimates of DNA concentration were assessed using variance components analyses.

Results

The OD method was the DNA quantification method most concordant with the reference sample among the three methods evaluated. A large fraction of the total variance for all three methods (36.0-95.7%) was explained by sample-to-sample variation, whereas the amount of variance attributable to sample handling was small (0.8-17.5%). Residual error (3.2-59.4%), corresponding to un-modelled factors, contributed a greater extent to the total variation than the sample handling procedures.

Conclusion

The application of a specific DNA quantification method to a particular molecular genetic laboratory protocol must take into account the accuracy and precision of the specific method, as well as the requirements of the experimental workflow with respect to sample volumes and throughput. While OD was the most concordant and precise DNA quantification method in this study, the information provided by the quantitative PCR assay regarding the suitability of DNA samples for PCR may be an essential factor for some protocols, despite the decreased concordance and precision of this method.

Background

  Molecular genetic studies hold the promise of identifying genetic factors that influence human disease susceptibility and outcome [1]. However, large sample sizes are required to find small to moderate sized genetic effects in association studies [2]. Technical advances in genotype and sequence analysis, together with essentially unlimited availability of sequence and sequence variation information, have substantially increased the scope of human genetic studies. It is now possible to examine many polymorphisms or sequences in high-throughput settings, using significantly less DNA per assay (~1-5 ng) than in the past decade. Nevertheless, conservation of precious DNA samples represents a critical goal for the efficient utilization of research resources. One area of technological development essential for success in high-throughput genotyping is the quantification of DNA. Accurate and

precise DNA quantification is necessary for efficient high-throughput genotyping and sample conservation. Inaccuracy in quantification of DNA results in the unnecessary consumption of DNA. Imprecise quantification increases variability in the amount of PCR product used by most genotyping technologies, leading to lower confidence in scoring of genotypes [3,4]. Inaccuracies in quantification of DNA samples that will constitute a pool of DNA lead to inaccurate allele frequency estimations [5]. Conservation of the original DNA samples is important to validate previous findings and to allow for future studies, even if whole genome amplification technologies prove to be accurate and reliable [6,7].

Until recently, UV absorbance spectroscopy has been the traditional method of quantifying DNA in molecular biology laboratories prior to molecular genetic analyses [8,9]. Quantification of genomic DNA using intercalating fluorochromes [10,11] and oligonucleotide hybridization methods [12] has decreased DNA consumption due to the increased sensitivity of these methods and has increased laboratory efficiency due to the high-throughput format of fluorometers. We assessed the performance of three DNA quantification methods and give recommendations for their use in facilities processing biospecimens for genotyping and sequencing analyses. These three methods are: 1) UV absorbance spectroscopy, referred to here as the OD method, 2) the detection of fluorescent signal enhancement of PicoGreen® dsDNA Quantitation Reagent which selectively binds to dsDNA [10,13], referred to here as the PG method, and 3) the detection of fluorescent signal from a 5' exonuclease assay [14], referred to here as the quantitative genomic or QG method.

First, we performed a high-throughput laboratory sample handling workflow using 22 high-quality DNA samples at concentrations typical of DNA extracted from lymphocytes. Sample-handling procedures diluted the DNA to two concentrations commonly used for normalization prior to genotyping. Three methods were used to estimate the DNA concentration of the aliquots producing a total of 14,784 concentration estimates. Variance components analysis was used to assess the influence of these sample-handling steps on performance of the three DNA quantification methods.

Results

Table 1 shows the summary statistics for the concentration estimates and Figure 1 presents the distributions of DNA concentration estimates for the OD-U, OD-D, QG and PG methods. The mean DNA concentration estimate for the OD-U evaluation, 20.57 ng/uL, was significantly higher than the expected value of 20.00 ng/uL (p-value for the t-test for the mean <0.0001); however, the distribution of the concentration estimates was negatively skewed (skewness coefficient = -0.12) due to several DNA concentration estimates in the 17-18 ng/uL range (Table 1 and Figure 1a). The expected DNA concentration estimate for OD-D, QG, and PG is 5.14 ng/uL, one-fourth of the observed OD-U estimate of 20.57 ng/uL, due to the dilution step (See Figure 2). The mean OD-D concentration estimate of 5.13 ng/uL was close to the expected value (p<0.5464), while the distribution of the concentration estimates was positively skewed (skewness coefficient = 0.25) (Table 1 and Figure 1b). The observed mean for QG, 5.98 ng/uL, showed the largest deviation from the expected value, (p<0.0001), and the distribution of the values was negatively skewed (skewness coefficient = -0.75), due primarily to a number of QG reactions (0.2% of the total) that failed to amplify and gave zero values (Table 1 and Figure 1c). The PG method also resulted in a mean concentration, 5.21 ng/uL, that was significantly

greater than the expected value (p<0.0148), and the concentration estimates were also negatively skewed (skewness coefficient = -0.52) due to several low, but non-zero concentration estimates (Table 1 and Figure 1d).

The CVs and their 95% CIs were low for all methods (Table 1). The CV for the OD-U method, 8.0%, was statistically significantly lower than the CVs for the other three methods, and the CV for the OD-D method, 10.3%, was statistically significantly lower than the CVs for the QG (13.8%) and PG (13.6%) concentration estimates, which were not statistically significantly different from each other.

Variance contributions of the individual samples, laboratory sample handling for each of the three DNA quantification methods were obtained using variance components analysis. The results are given in Table 2. Overall, the OD methods had the smallest total variance, and the fluorometric methods had the largest. The percent error attributable to sample variability was greatest for the 20 ng/uL dilution of DNA in the OD-U method (95.7%), and much less for the methods that had the additional 5 ng/uL dilution of DNA (36.0-60.0%), though the absolute amount of error attributable to sample variability was significantly greater for the fluorometric methods than for the OD methods. The contribution to the total variance from assay aliquot or dilution was less than 5% for all three methods. The variance contribution of OD Aliquot was substantial for OD-D (12.8%), but less than 3% for all the other methods. The error due to repeated OD readings of the same DNA sample in the same well was estimated to be zero. The amount of variance attributable to the entire laboratory sample handling procedure was largest in the OD-D method (17.5%) and smallest in the PG method (4.3%). The variance attributable to the OD Aliquot step in the OD-D method (12.8%) was significantly larger than any other error component attributable to laboratory sample handling procedures. Residual error accounted for a larger proportion of variance in the fluorometric methods (36-59.4%) than in the OD methods (3.2-22.8%).

Discussion

Accurate and precise quantification of DNA is critical for efficient genotyping, particularly in a high-throughput setting, in order to maximize completion rates, accuracy and reproducibility. Accurate quantification of DNA enables prudent management of DNA samples. Precise quantification of DNA reduces the variance of fluorescent signal derived from high-throughput genotyping technologies and increases the confidence in scoring genotypes. We have compared the performance of three different DNA quantification methods widely used in biorepository and genotyping facilities. The optical density of DNA (260/280 nm), the most popular approach to estimate DNA concentrations, was assessed at two different DNA concentrations (20 ng/uL and 5 ng/uL) commonly used for normalization of DNA prior to genotyping workflows. The fluorometric methods that we assessed were PicoGreen®, which fluoresces when bound to dsDNA, and the QG method, based on the TaqMan® assay. The OD-D (5 ng/uL) method was the least biased method, and the OD-U (20 ng/uL) method was the most precise method. The QG method was the most biased method, and both fluorometric methods were less precise than the OD methods evaluated. Thus, absorbance spectrophotometry of DNA appears to be less biased than the more complex fluorometric methods that rely upon multiple interacting components and more indirect measurements of DNA concentration. However, the experimental design used to test the methods may have influenced the apparent relationship between the complexity and the performance of the method.

Specifically, the more complex fluorometric methods were evaluated after a larger number of laboratory sample handling steps and the volumes used are inversely related to the complexity of the DNA quantification assays.

Sample-to-sample variability was the single largest factor that contributed to the error in estimated DNA concentrations, while laboratory sample handling steps and residual (unmodeled) factors had a lesser impact, as assessed by variance components (Table 2). DNA sample preparation can be difficult to control in some large-scale studies; thus, genotyping facilities receiving DNA samples from multiple studies should expect DNA sample variability to be a major factor affecting sample-processing throughput. Reassuringly, the estimated residual variance was minimal in the quantification method with the smallest number of sample handling steps (OD-U). The residual variance was greatest in the fluorometric methods, perhaps because the greater complexity and reduced reproducibility of these assays. The relationship between aliquot volume and sample variability also suggests that smaller volumes, e.g., the 2.0 uL and 5.0 uL of DNA sample used in the QG and PG methods, respectively, introduce more sample variability, perhaps due to sampling effects or increased failure to transfer small volumes.

Spectrophotometric determination of DNA concentration provided at high concentration by the biorepository, using one or two dilution steps, each step associated with the estimation of concentration via OD, was the least biased and exhibited the smallest total variance in this study. Thus, quantification workflows that consist of only one OD measurement of one aliquot after DNA extraction and then aliquoting or normalization to the nominal concentration used in a genotyping or sequencing assay are likely to be significantly more inaccurate than the protocol followed in this study. Other advantages of the OD method are that spectrophotometers are more widely available than the fluorometric laboratory equipment required for the QG and PG methods and that the OD estimation of DNA concentration can be performed without consuming the sample, additional reagents or incubation time. In contrast, the QG and PG methods require more specialized optical equipment, consume DNA samples, and are dependent upon reagents and a reaction timeline that is measured in minutes to hours. QG requires a labeled oligonucleotide, *Taq* DNA polymerase, deoxynucleotides, buffer, a sophisticated detection instrument, and ~ two hours. The additional complexity of the QG and PG methods suggest that the OD method is preferred over the QG and PG methods.

Despite these advantages of the OD method, there are disadvantages compared to the fluorometric methods. Large amounts of DNA are necessary for spectrophotometric analysis in current instrumentation for high-throughput environments. In the OD methods used in this study, the minimum amount and volume of DNA evaluated was 750 ng in 150 uL. Thus, fluorometric determination of DNA concentration may be preferred over the OD method in high-throughput environments principally due to the lower quantity of DNA needed. Proteins, RNA, and salts, all of which are contaminants of DNA extracted from various biological sources, can increase the spectrophotometric estimation of DNA concentration. In addition, the OD method cannot distinguish between ssDNA and dsDNA in solution as can the PG method, which does not bind to ssDNA [13] or between DNA from the species of interest (in our study human DNA) and other potentially contaminating sources of DNA, for example, bacterial DNA, as can the QG method. The ability of a DNA sample to successfully amplify via PCR is an indication of the quality of the DNA sample, and this information is provided only by the QG method. Thus, the OD method has limitations that may contribute to inaccuracy of DNA concentration

estimates and it may not be sufficiently sensitive for some high-throughput environments, especially those working with limited or potentially contaminated DNA samples, such as forensic samples or some population-based collections.

Rengarajan et al. have evaluated the sensitivity and linearity of four fluorochromes (ethidium bromide, Hoechst 33258, SYBR® Green I and PicoGreen®) commonly used for DNA quantification [15]. They demonstrate that ethidium bromide and Hoechst 33258 are suitable for estimating DNA concentration in the range of ~200 ng/mL to 20 ug/mL and that SYBR® Green I and PicoGreen® are suitable for estimating DNA concentration in the range of 0.2 ng/mL to 20 ug/mL. The intra-assay CV (between replicates) of PicoGreen® was stated to be 8.3%, where the intra-assay CV of PicoGreen® in this study was found to be 13.6%, which is comparable to their findings. In this study, however, we are evaluating the bias of DNA quantification methods, not the sensitivity of DNA quantification methods.

Real-time PCR assays are commonly used to detect small changes in gene copy-number and expression, and the reference values used in those applications are often in the femtogram and picograms ranges [16], and use standard curves based on a log-scale. In this study, however, the QG method was used to determine quantity of DNA, and the 5.0 ng/uL concentration was chosen to be indicative of values commonly seen in high-throughput laboratories, such as a biorepository. Therefore, a linear standard curve range (1.0 to 10.0 ng/uL) was chosen, and the assay's repeatability was high, with the CV of the R^2-value of the standard curve for QG method only 0.22%, and only 0.14% for the PG method. Although SYBR® Green I has been used to detect dsDNA product during PCR [17], it has been previously shown that the greatest specificity is achieved with probe hybridization [18], as used in this study with the QG method. Although the use of SYBR® Green I is also suitable for dsDNA quantification, Molecular Probes Inc., the manufacturer of both dyes, claims that PicoGreen® is more linear than SYBR® Green I in the range of 10 pg/uL to 1 ug/uL (personal communication). This linearity had been addressed in Renegarajan, et al., with the linearity of PicoGreen® (0.9737) slightly higher than for SYBR® Green I (0.9623).

The insensitivity of the OD and PG methods to potential contaminants of DNA solutions and the sequence specificity and functional information that the QG method provides may outweigh the increased inaccuracy and reduced precision seen in this analysis of the QG method for some practical applications. The choice of DNA quantification methods to be used by a particular laboratory should be guided by an understanding of the accuracy, precision and potential sources of error of those methods, and the practical aspects of implementing a particular DNA quantification method for a particular DNA sample processing or genotyping application.

Conclusions

We have evaluated bias, variance, and variance components associated with three common methods of DNA quantification in a high-throughput setting. Although the spectrophotometric method was the most accurate and precise quantification method evaluated in this study, the functional information obtained from the real-time PCR assay provides human specific data on the quantity and quality of the DNA sample. Sample-to-sample variability accounted for the majority of quantification error for all three methods, whereas the variance associated with sample handling procedures was minimal. High-throughput laboratories and biorepositories should be aware of the

sources of error in sample handling procedures and quantification methods to ensure high completion and low inaccuracy rates in genotyping workflows.

Methods

DNA samples

  22 DNA samples were obtained from Coriell Cell Repositories (Camden, NJ, USA). These samples were extracted from cell lines using standard high salt procedures [19] and resuspended by the repository in 10 mM Tris-1mM EDTA, pH=8.0. The samples were provided at high concentrations (Mean: 350.0 ng/uL; σ: 28.6 ng/uL; Range: 244.0 – 375.0 ng/uL), as determined by the repository using optical density (OD_{260} and OD_{280}).

Instruments

  OD measurements were obtained using a SpectraMax Gemini Plus spectrophotometer, hereafter "Spectrophotometer" (Molecular Devices Corp., Sunnyvale, CA, USA). The spectrophotometer was calibrated at the Core Genotyping Facility (CGF) according to manufacturer's protocols using the SPECTRAtest Validation Package for Microplate Spectrophotometers (Molecular Devices Corp., Sunnyvale, CA, USA). This validation plate contains a series of National Institutes of Standards and Technology (NIST)-traceable filters.   Eight aspects of the spectrophotometer were calibrated: 1) baseline noise, 2) ultimate dark, 3) optical alignment, 4) absorbance accuracy, 5) absorbance precision, 6) stray light, 7) wavelength accuracy, and 8) wavelength precision. QG was performed on an ABI 7900HT Sequence Detector System, hereafter "ABI 7900" (Applied Biosystems, Inc., Foster City, CA, USA). The ABI 7900 was calibrated at the CGF according to manufacturer's protocol using the Sequence Detection Systems 384-well Spectral Calibration Kit (Applied Biosystems, Inc., Foster City, CA, USA). Enhancement of PicoGreen® fluorescence was measured on a Fluoroskan Acent FL, hereafter "Fluorometer" (Thermo LabSystems, Franklin, MA, USA). The fluorometer was calibrated at the CGF according to the manufacturer's protocol. Addition of QG reaction mix was done using a Multimek 96 Automated 96-channel pipettor, hereafter "Multimek 96" (Beckman Coulter, Fullerton, CA, USA). Aliquots for QG and PG assays were dispensed using a Hydra 96 liquid handling system, hereafter "Hydra" (Robbins Scientific Corp., Sunnyvale, CA, USA). All aliquots, dilutions, and assay plates (Figure 2) were prepared in 96-well plates, with the exception of QG, which was performed in 384-well plates. All liquid handling systems and manual pipettes were calibrated according to the manufacturer's recommendations.

Optical Density Assay (OD)

  The Optical Density assay (OD) was tested at two concentrations, 20 ng/uL and 5 ng/uL (Figure 2). According to the biorepository's supplied OD estimation of DNA concentration, a 20.0 ug aliquot of each DNA sample was manually diluted to a nominal 20 ng/uL (1000 uL). Four 150.0 uL "OD Undiluted" aliquots (OD-U) at ~ 20 ng/uL (~ 3.0 ug) were removed and the OD (260, 280, and 325 nm) was determined using a Spectrophotometer. Eight OD readings ("Run") were obtained for each of the four aliquots for each sample, resulting in a total of 704 data points. From

each of the four OD-U aliquots, four 37.5 uL "OD Diluted" aliquots (OD-D) were manually removed and diluted 1:4 to ~ 5 ng/uL (~ 750 ng) using the OD readings determined in this laboratory at the "Run" step. Eight OD readings were obtained from each OD-D dilution, for a total of 2816 data points. To calculate DNA concentration, OD absorbance values ($A°_{260}$) were multiplied by 50, assuming 1.0 A° = 50.0 ng/uL dsDNA.

Quantitative Genomic PCR Assay (QG)

A real-time PCR assay was developed to quantify human genomic DNA using a non-polymorphic region of the human BRCA1 locus. Primers were empirically chosen by testing three primer pairs for each probe with a serial dilution of human genomic DNA, quantified multiple times by OD spectrophotometry, to determine the combination that resulted in the highest R2 value (typically > 0.99) when plotting the logarithm (input DNA) versus the threshold cycle. The best primer pair/probe combination was then tested against 100 ng of genomic DNA from *S. mutans*, *S. cerevesiae*, mouse, dog, chicken, cat, and bovine as input, none of which gave signal that was above background after 40 cycles. According to BLAST [20], the amplicon was found not to be similar to any regions other than *BRCA1* [3×10^{-60} versus 0.33 as the next lowest Expect (E) value in the human genome (Build 30)]. A 119 bp region at the *BRCA1* region (IVS11+126) was amplified using the forward primer 5'-AAACATGTTCCTCCTAAGGTGCTTT-3', and the reverse primer 5'-ATGAAACCAGAAGTAAGTCCACCAGT-3'. Detection of this amplicon is achieved by a 41-bp labeled oligonucleotide probe, 5'-FAM-CCTTCACACAGCTAGGACGTCATCTTT-TAM-3'.
From each of the 16 OD-D aliquots, 16 2.0 uL "Assay Aliquots" (~10 ng) were robotically transferred using a Hydra to 384-well plates and subjected to quantification using the QG method (Figure 2 and Standard DNA Preparation). Prior to QG, water was removed from the 2.0 uL DNA "Assay Aliquots" overnight in a desiccator. Reaction conditions for QG are as follows: 400 nM primers, 50 nM probe, 1X Universal Master Mix (Applied Biosystems Inc., Foster City, CA, USA), q.s. with water to a final reaction volume of 5.0 uL. Reactions were cycled on ABI 7900 with the following cycle conditions: 50 °C for 2', 95 °C for 10', followed by 40 cycles of 95 °C for 30" and 60 °C for 1'. QG reactions were cycled using the "Absolute Quantification" assay setting on the instrument. Data was analyzed using Sequence Detection Software 2.0 (Applied Biosystems Inc., Foster City, CA, USA), with "Automatic C_t (Cycle Threshold) Determination" enabled with a minimum quality score of 5 (out of 10).

PicoGreen® Assay (PG)

From the each of the OD-D aliquots, 16 5.0 uL "Assay Aliquots" (~ 25 ng) were robotically transferred using a Hydra to 96-well microplates and subjected to quantification using the PG method (Figure 2 and Standard DNA Preparation). Fluorescence of PicoGreen® was detected using a Fluorometer, where 95.0 uL of 1X PicoGreen® Reagent, diluted in 1X TE Buffer, was robotically dispensed by the Fluorometer into 5.0 uL of sample. Samples were incubated in darkness for 5 minutes, and then excited (485 nm) and emission measured (520 nm).

Standard DNA Preparation

  QG and PG methods of DNA quantification require the use of a standard curve derived from DNA samples of known concentrations and an experimentally derived cycle threshold (C_t) value, for QG, and raw fluorescence unit (RFU) values, for PG. A DNA sample was obtained from Coriell Cell Repositories and used to generate a set of known DNA concentrations (Figure 3). According to the biorepository's concentration estimate, a 20.0 ug aliquot was diluted to 20.0 ng/uL, and the OD of 100.0 uL aliquots (N=10) were determined using a spectrophotometer ("OD-STD-20"; Mean: 21.10 ng/uL; σ: 1.33 ng/uL; Range: 19.20- 23.26 ng/uL). Based on the measured OD_{260} values, each of the ten aliquots was diluted with water to 12.0 ng/uL, combined into a single solution and homogenized by vortexing. 100.0 uL aliquots (N=19) were removed and OD determined ("OD-STD-12"; Mean: 11.65 ng/uL, σ: 0.49 ng/uL, Range: 10.90 – 12.84 ng/uL). According to the measured OD_{260} values, each of the 19 aliquots was further diluted with water to 10.0 ng/uL. All aliquots were combined into a single solution and homogenized by vortexing. 100.0 uL aliquots (N=21) were removed and OD determined ("OD-STD-10"; Mean: 10.041 ng/uL, σ: 0.46 ng/uL, Range: 9.43 – 11.08 ng/uL). All 21 aliquots were combined into a single solution and homogenized by vortexing.
  The "OD-STD-10" was used to generate a set of standard solutions at 1.0, 2.0, 4.0, 5.0, 6.0, 8.0, and 10.0 ng/uL for QG and PG quantification. Four aliquots from each of the seven standard solutions comprised a single standard curve consisting of 28 data points. For every four QG and PG assay aliquots quantified, one standard curve was prepared and used to determine DNA concentration values. In total, 16 standard curves were generated for each method (total N=32), resulting in 448 data points collected from the standard curves for each method (total N=896). Mean R^2-value for QG was 0.995 (σ: 0.002, CV: 0.22%), and 0.998 (σ: 0.001, CV: 0.14%) for PG. The mean C_t value for the QG method was 0.0987 (σ: 0.016, CV: 16.21%). For QG, the slope of the standard curve is indicative of the efficiency of the assay. At 100% efficiency, a reaction should achieve a slope of –3.33, since every 10-fold difference in quantity translates to a difference on 3.33 C_t's (ABI 7900HT Sequence Detection System, User Guide). The mean slope for QG was –3.836 (σ: 0.087).

Database Design

  In order to track the multiple aliquots and dilutions of each DNA sample used in this study, a Microsoft® Access database was created. The database schema allows the import of concentration estimates generated from each of the quantification platforms. The plate name and well position provided in the instruments' output file identified the sample and from which dilutions and/or aliquots it originated. For the QG and PG methods, the database also contained the raw values (C_t and RFU) of each of the standard DNA data points used in determining the DNA concentration.

Statistical Analysis

  We calculated descriptive statistics for the measurements, including means, standard deviations, the coefficient of variation (CV) and its 95% confidence interval (CI), and tested if the mean amount for each method was equal to its expected value using t-tests. We estimated the components of variance attributable to various steps in the sample handling for each method. For the variance components analysis,

measurements were analyzed on the natural logarithmic scale, to reduce the dependence of the SD of the response in the mean. This allowed the assumption that the total variance of the mean was unrelated to individual sample concentration (see Cook and Weisberg, "Applied Regression Including Computing and Graphics" Wiley Series in Probability and Statistics. 1999, page 317). For each method, a nested component of variance analysis was performed using the restricted maximum likelihood approach (PROC VARCOMP, SAS 8.0, SAS/Stat User's Guide, Version 8. SAS Institute Inc.: Cary, NC, USA, 1999). Restricted maximum likelihood estimates agree with the usual analysis of variance (ANOVA) estimates when all estimates are greater than zero. Components were estimated for "Sample", "OD Aliquot", "Dilution", "Assay Aliquot", and "Run" (Figure 2). The general statistical model was written:

$$\text{Log}(z_{ijkl}) = \mu + a_i + b_{j(i)} + c_{k(ij)} + d_{l(ijk)} + \varepsilon_{m(ijkl)},$$

with z_{ijkl} denoting the amount of DNA for sample i (i=1, ..., 22). In this model, μ was the mean DNA concentration and a_i, $b_{j(i)}$, $c_{k(ij)}$, $d_{l(ijk)}$ and $\varepsilon_{m(ijkl)}$ were normally distributed random variables with mean zero and variances σ^2_a, σ^2_b, σ^2_c, σ^2_d and σ^2_e, respectively. For OD-U, the model reduced to:

$$\text{Log}(z_{ijkl}) = \mu + a_i + b_{j(i)} + c_{k(ij)} + \varepsilon_{l(ijk)},$$

with the variance component σ^2_a corresponding to "Sample", σ^2_b to "OD Aliquot" and σ^2_c to "Run". For PG and QG, the variance component σ^2_a corresponded to "Sample", σ^2_b to "OD Aliquot", σ^2_c to "Dilution", and σ^2_d to "Assay Aliquot". For OD-D, σ^2_a corresponded to "Sample", σ^2_b to "OD Aliquot", σ^2_c to "Dilution" and σ^2_d to "Run". For all methods, σ^2_e represented all sources of variance that are not specified in the model.

Authors' contributions

KH participated in the design of the study, performed the wet laboratory experiments, and assisted in preparing the manuscript. RP performed the statistical analysis on the DNA quantification data collected and assisted in preparing the manuscript. MB created and managed the database of DNA quantification estimates. JS designed the QG-PCR assay. SC assisted in the preparation of the manuscript. AWB participated in the design of the experiment and assisted in the preparation of the manuscript. All authors read and approved the final version of the manuscript.

Acknowledgements

References

1. Risch N, Merikangas, K: **The future of genetic studies of complex human diseases.** *Science* 1996, **273:**1516-17.
2. Camp NJ: **Genomewide transmission/disequilibrium testing-- consideration of the genotypic relative risks at disease loci.** *Am J Hum Genet* 1997, **61:**1424- 30.
3. Duewer DL, Kline MC; Redman JW, Newall P J, Reeder DJ: **NIST mixed stain studies #1 and #2: interlaboratory comparison of DNA quantification practice and short tandem repeat multiplex performance with multiple-source samples.** *J Forensic Sci* 2001, **46:**1199-1210.
4. Kline MC, Duewer DL, Redman JW, Butler JM: **NIST mixed stain study 3: DNA quantitation accuracy and its influence on short tandem repeat multiplex signal intensity** *Anal Chem* 2003, **75:**2463-69.
5. Sham P, Bader JS, Craig L, O'Donovan M, Owen M: **DNA pooling: a tool for large scale association studies.** *Nat Rev Genet* 2002, **3:**862-871.
6. Dean FB, Hosono S, Fang L, Wu X, Faruqi AF, Bray-Ward P, Sun Z, Zong, Q, Du Y, Du J, Driscoll M, Song W, Kingsmore SF, Egholm M, Lasken RS: **Comprehensive human genome amplification using multiple displacement amplification.** *Proc Natl Acad Sci USA* 2002, **99:**5261-66.
7. Hosono S, Faruqi AF, Dean FB, Du Y, Sun Z, Wu X, Du J, Kingsmore SF, Egholm M, Lasken RS: **Unbiased whole-genome amplification directly from clinical samples.** *Genome Research* 2003, **13:**954-64.
8. Beaven GH, Holiday ER, Johnson EA: In *The Nucleic Acids, Vol. 1,* **1955,** Chargaff and Davidson JN. (ed.). Academic Press, New York.
9. McGown EL: **UV absorbance measurements of DNA in microplates.** *Biotechniques* 2000, **60:**63-4.
10. Ahn SJ, Costa J, Emanuel JR: **PicoGreen quantitation of DNA: effective evaluation of samples pre- or post-PCR.** *Nucleic Acids Res* 1996, **24:**2623-25.
11. Vitzthum F, Geiger G, Bisswanger H, Brunner H, Bernhagen J: **A quantitative fluorescence-based microplate assay for the determination of double-stranded DNA using SYBR Green I and a standard ultraviolet transilluminator gel imaging system.** *Anal Biochem* 1999, **276:**59-64.
12. Sanchez JL, Storch GA: **Multiplex, quantitative, real-time PCR assay for cytomegalovirus and human DNA.** *J Clin Microbiol* 2002, **40:** 2381-2386.
13. Singer VL, Jones LJ, Yue ST, Haugland RP: **Characterization of PicoGreen reagent and development of a fluorescence-based solution assay for double-stranded DNA quantitation.** *Anal Biochem* 1997, **249:**228-38.
14. Heid CA, Stevens J, Livak KJ, Williams PM: **Real time quantitative PCR.** *Genome Res* 1996, **6:**986-94.
15. Rengarajan L, Cristol SM, Mehta M, Nickerson JM: **Quantifying DNA concentrations using fluorometry: a comparison of flourophores.** *Mol Vis* 2002, **8:**416-421.
16. Sofi Ibrahim M, Kulesh DA, Saleh SS, Damon IK, Esposito JJ, Schmaljohn AL, Jahrling PB: **Real time PCR assay to detect smallpox virus.** *J Clin Microbiol* 2003, **41:**3835 -3839.
17. Stagliano KE, Carchman E, Deb S: **Real-time polymerase chain reaction quantitation of relative expression of genes modulated by p53 using SYBR Green l.** *Methods Mol Biol* 2003, **234:**73-91.

18. Newby DT, Hadfield TL, Roberto FF: **Real-time PCR detection of Brucella abortus: a comparative study of SYBR green l, 5'-exonuclease, and hybridization probe assays**. *Appl Environ Microbiol* 2003, **69:**4753-4759.
19. Miller SA, Dykes DD, Polesky HF: **A simple salting out procedure for extracting DNA from human nucleated cells.** *Nucleic Acids Res* 1988, **16:**1215.
20. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: **Basic local alignment search tool.** *J Mol Biol* 1990, **215:**403-10.

Figures

Figure 1.　　Distribution of DNA concentration estimates [ng/uL] for the three quantification methods (four evaluations). The estimates are standardized by dividing the counts by the total number of observations for each of the four evaluations. Figure 3a. Distribution of DNA concentration estimates [20 ng/uL] for OD-U (N=704 observations). Figure 3b, 3c, 3d. Distribution of DNA concentration estimates [5 ng/uL] for OD-D (N=2816), QG (N=5632), and PG (N=5632), respectively.

Figure 2.　　Experimental Workflow. Flowchart shows sample-handling processes (dilutions and aliquots, depicted by arrows) for a single DNA sample (right). Total DNA concentration estimates collected for all samples (N=22) for each method (four evaluations) is shown at left.

Figure 3.　　Creation of a standard set of DNA solutions. OD methods were used to determine the DNA concentration of a standard DNA solution that was subsequently used for QG and PG quantification methods.

Tables

Table 1.　　Descriptive statistics for the three DNA quantification methods (four evaluations).

Evaluation Type	N	Obs. Mean [ng/uL]	Std. Dev. [ng/uL]	Exp. Mean [ng/uL]	*P*-value	CV (%) (95% CI)	Min. Value [ng/uL]	Max. Value [ng/uL]	Skewness
OD-U	704	20.57	1.64	20.00	< 0.0001*	8.0 (0.079, 0.080)	17.17	24.41	- 0.12
OD-D	2816	5.13	0.53	5.14	0.5464	10.3 (0.097, 0.109)	3.63	8.26	0.25
QG	5632	5.98	0.83	5.14	< 0.0001*	13.8 (0.129, 0.146)	0.00	11.41	- 0.75
PG	5632	5.21	0.71	5.14	0.0148*	13.6 (0.129, 0.143)	0.19	7.95	- 0.59

Table 2. Variance component estimates (standard error in parenthesis) for the three DNA quantification methods (four evaluations).

| Variance Component	N (Replicates)	OD-U	OD-D	QG	PG
Sample	22	0.0064 (0.002)	0.0067 (0.0021)	0.0110 (0.0034)	0.0130 (0.004)
% of Total Variance		95.7%	60.0%	55.0%	36.0%
OD Aliquot	4	0.0000 (0)	0.0014 (0.0013)	0.0005 (0.0005)	0 (0)
% of Total Variance		0.8%	12.8%	2.3%	0.0%
Dilution	4	N/A	0.0005 (0.0002)	0.0004 (0.0002)	0.0011 (0.0004)
% of Total Variance	Nested w/ OD Aliq.		4.7%	1.8%	3.2%
Assay Aliquot	16	N/A	N/A	0.0010 (0.0001)	0.0004 (0.0001)
% of Total Variance				4.8%	1.1%
Run	8	0.0000 (0)	0 (0)	N/A	N/A
% of Total Variance		0.0%	0.0%		
Residual Error	-	0.0002 (0)	0.0025 (0.0001)	0.0072 (0.0001)	0.0212 (0.0004)
% of Total Variance		3.2%	22.8%	36.0%	59.4%
Total Variance	-	0.0066	0.0111	0.0201	0.0357
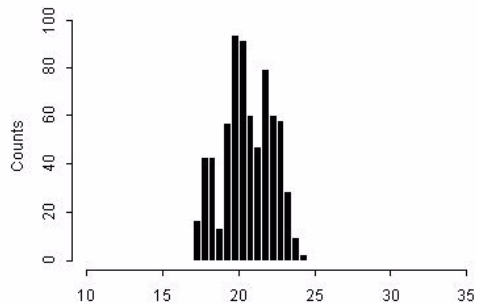
Fig 1a: DNA concentrations for OD-U
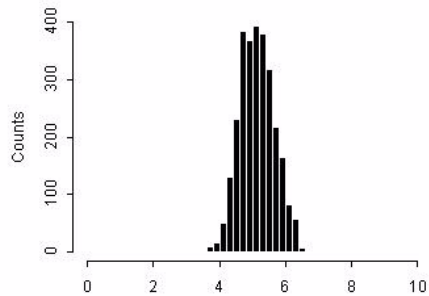
Fig 1b: DNA concentrations for OD-D
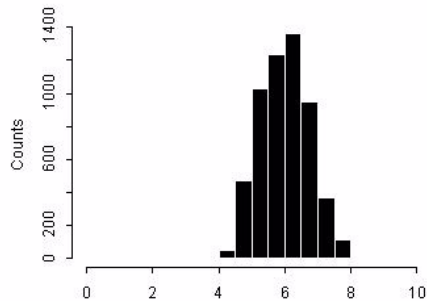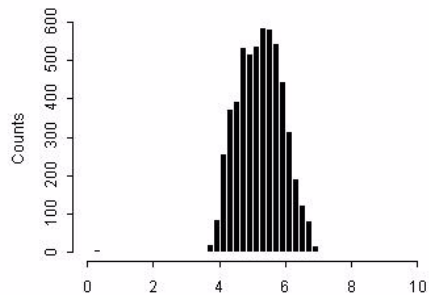
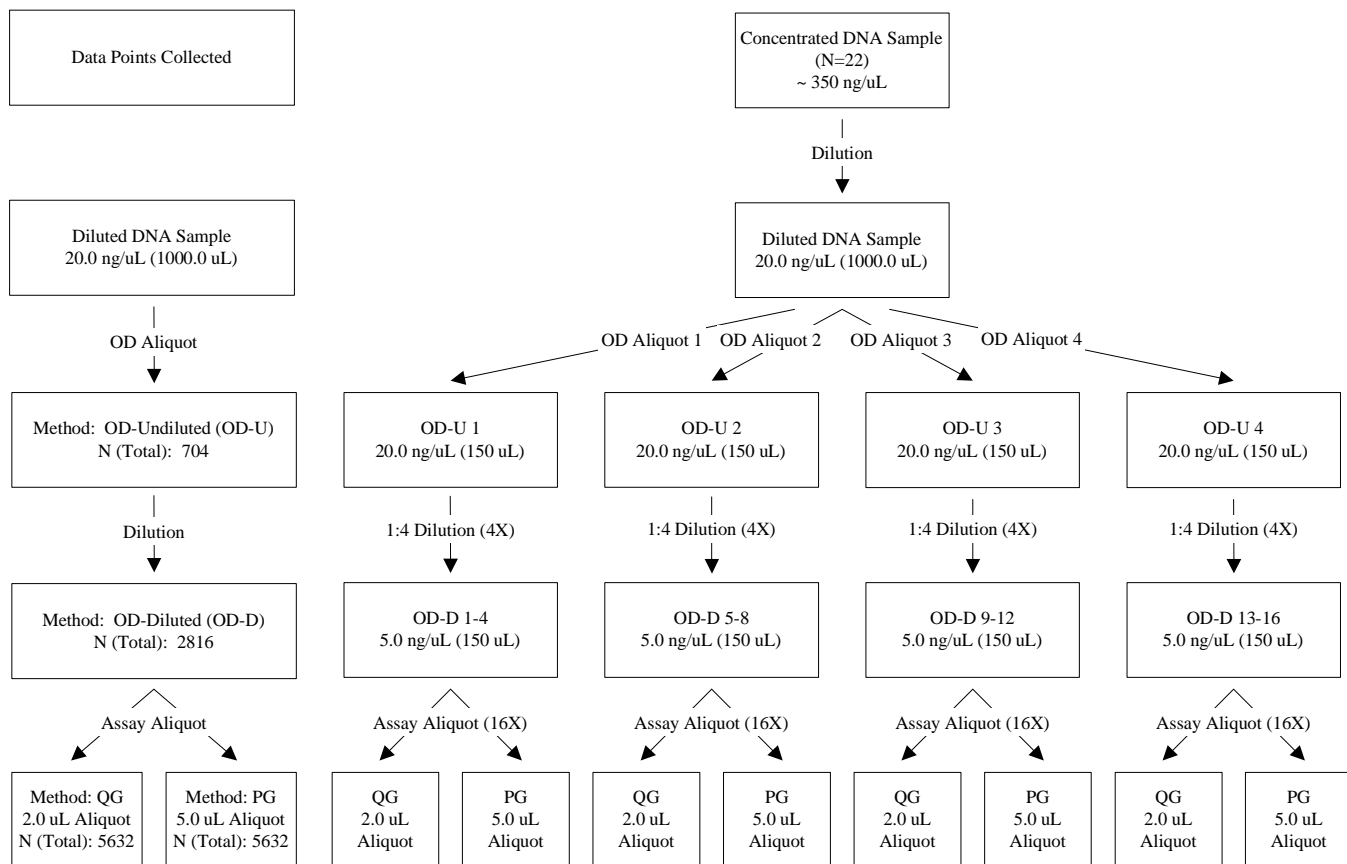Fig 1c: DNA concentrations for QG

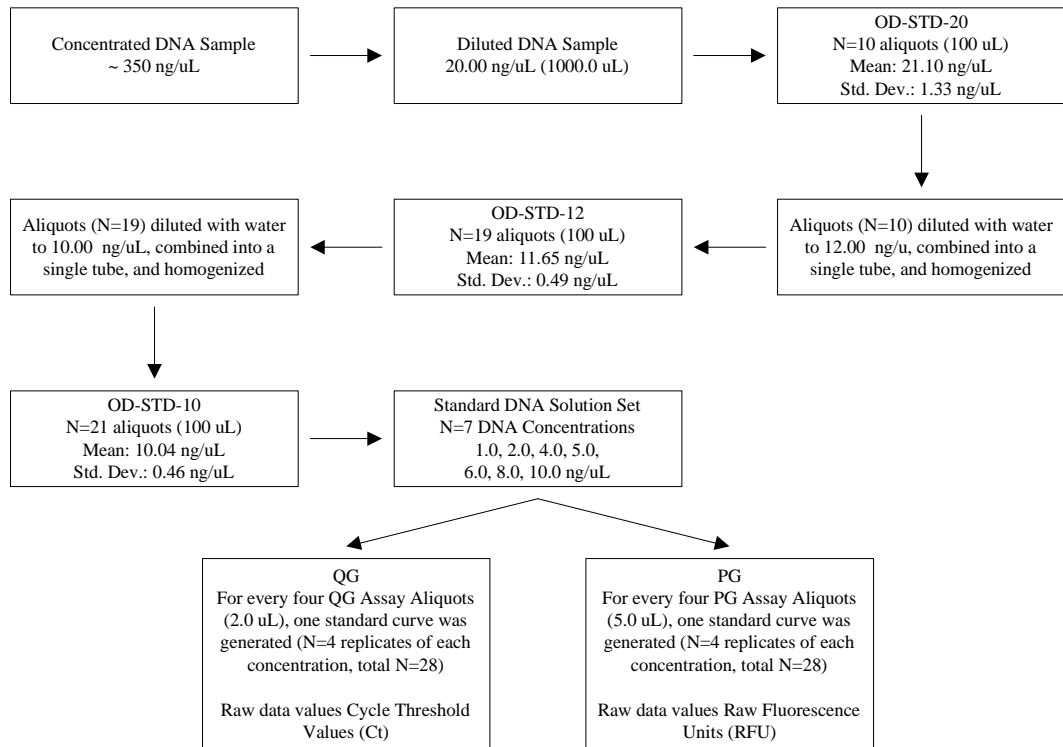Fig 1d: DNA concentrations for PG

Figure 1

Figure 2

Figure 3